

Overview of the ATLAS distributed computing system

Johannes Elmsheuser^{1,*}, and Alessandro Di Girolamo² for the ATLAS collaboration

¹Brookhaven National Laboratory, Upton, NY, USA

²CERN, Geneva, Switzerland

Abstract. The CERN ATLAS experiment successfully uses a worldwide computing infrastructure to support the physics program during LHC Run 2. The Grid workflow system PanDA routinely manages 250 to 500 thousand concurrently running production and analysis jobs to process simulation and detector data. In total more than 370 PB of data is distributed over more than 150 sites in the WLCG and handled by the ATLAS data management system Rucio. To prepare for the ever growing LHC luminosity in future runs new developments are underway to even more efficiently use opportunistic resources such as HPCs and utilize new technologies. This paper will review and explain the outline and the performance of the ATLAS distributed computing system and give an outlook to new workflow and data management ideas for the beginning of the LHC Run 3. It will be discussed that the ATLAS workflow and data management systems are robust, performant and can easily cope with the higher Run 2 LHC performance. There are presently no scaling issues and each subsystem is able to sustain the large loads.

1 Introduction

During Run 2 the Large Hadron Collider (LHC) continues to deliver large amounts of data to the experiments (see Fig. 1 left). The distributed computing system of the ATLAS experiment [1] as outlined in Fig. 1 right is built around the two main components: the workflow management system PanDA [3] and the data management system Rucio [4]. It manages the computing resources to process this data at the Tier-0 at CERN, reprocesses it once per year at the Tier-1 and Tier-2 WLCG [5] Grid sites and runs continuous Monte Carlo (MC) simulation and reconstruction. In addition continuous distributed analysis from several hundred ATLAS users is executed. The resources used are the Tier-0 at CERN and Tier-1/2/3 Grid sites worldwide and opportunistic resources at HPC (High Performance Computer) sites, Cloud computing providers and volunteer computing resources submitted via ATLAS@Home/Boinc [6]. All the infrastructure is monitored by different shift teams for production and distributed analysis. The ATLAS distributed computing system is performing very well in handling and processing all this data.

2 Data processing in 2017 and 2018

ATLAS has been using much more CPU resources than the officially pledged in 2017 and in 2018 so far (see Fig. 2 and Fig. 3). The pledge of CPU resources sums up to roughly 250k

*e-mail: johannes.elmsheuser@cern.ch

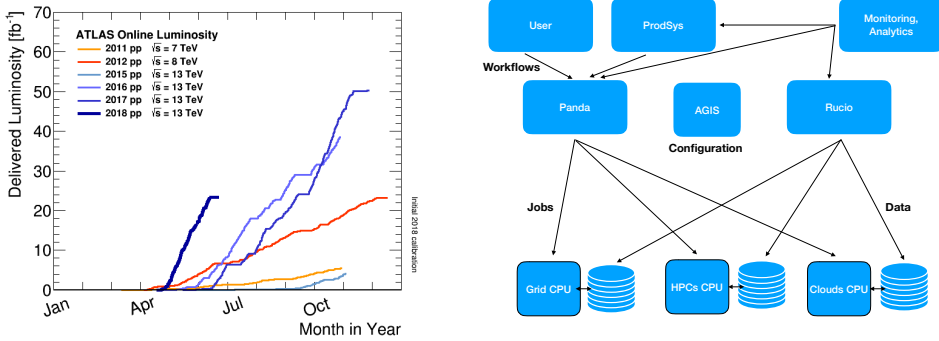


Figure 1. Left: ATLAS integrated luminosity during LHC Run-1 and Run 2 [2]. Right: Schema of the ATLAS distributed computing system and its components.

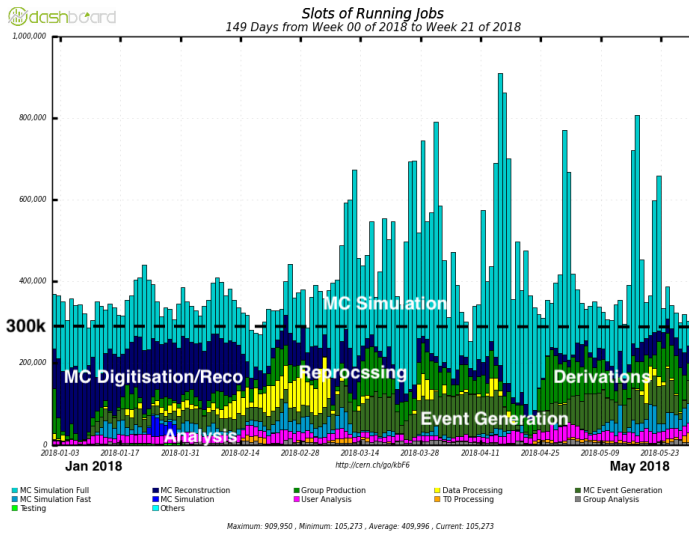


Figure 2. Number of running jobs by core count and workflow type in the first half of 2018.

concurrently running Grid job slots in 2017. Many ATLAS WLCG computing sites provide more CPU power over the pledge. In addition the ATLAS workflow management system is very efficiently exploiting additional non-WLCG opportunistic resources that have been mentioned in Sec. 1. Beyond pledge and opportunistic resources are mainly used for MC event generation and simulation and have allowed the production of almost twice the number of fully simulated events than would have been possible with the pledged resources only. Fig. 3 shows the normalised HEPSPC06 wall clock time which is calculated from average of the HEPSPC06 normalisation values of CPUs at a site and the number of running jobs at a site as shown in Fig. 2. CPUs at HPC centres are usually weaker than those at Grid sites and have lower HEPSPC06 normalisation values. This causes the peaks in the number of running jobs to disappear in the equivalent HEPSPC06 distribution.

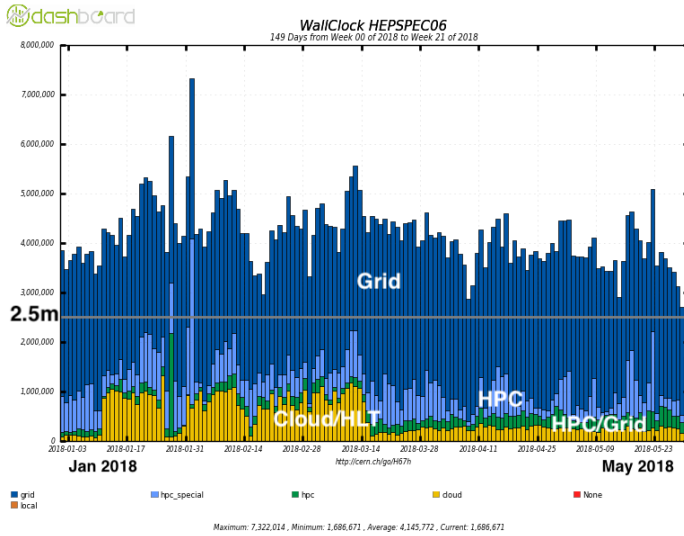


Figure 3. CPU usage normalised to HEPSP06 by resource type in the first half of 2018.

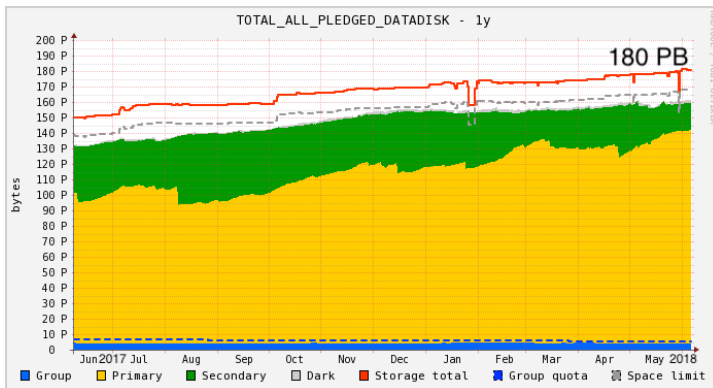


Figure 4. ATLAS data volume distribution over the last 12 months. Primary data (yellow) is partially replicated to several sites as well as cached data (green).

3 Data management

In the middle of 2018 the ATLAS data volume exceeds 370 PB in disk and tape storage. There are more than 20 million datasets with more than 1 billion files stored world-wide at all the ATLAS computing sites. About 180 PB of data are stored on disk and partially replicated with one additional copy at other sites to improve the availability and job throughput for popular data (see Fig. 4). The data transfers between the sites reach an aggregated traffic of 20 GB/s in a weekly average with daily averages up to 40 GB/s. On average 1.5-2 million files per day are transferred between different sites with a volume of more than 1 PB per day (see Fig. 5). A large fraction stems from delivering the input data to production jobs especially for the so called derivation framework. This workflow processes the reconstruction output AOD files into derived AOD files (DAOD) which are used in different subsequent physics

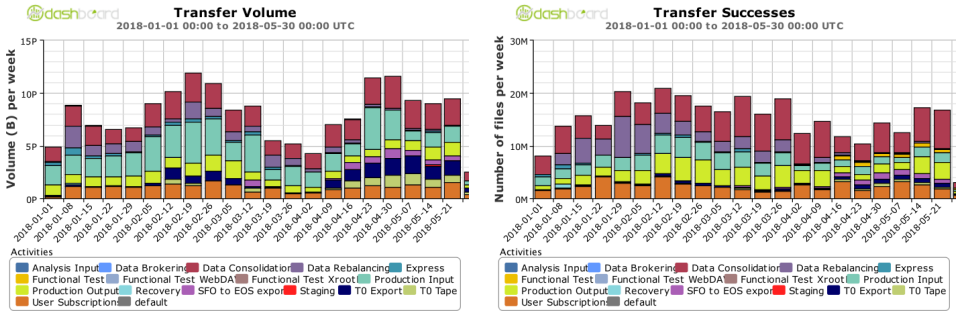


Figure 5. Left: Data transfers between the ATLAS sites aggregated by transfer activity and volume size. Right: The same information but aggregated by the number of files.

analysis. The data transfer system is currently limited by the maximal number of files that can be transferred simultaneously in the WLCG files transfer service FTS [7].

4 Production activities

The major production activities in 2017 and 2018 so far are the MC event generation, simulation and reconstruction campaigns MC15 for 2015 and 2016 data and MC16 for 2017 and 2018 data (see Fig. 6 left). These activities took about 72% of the Grid resources in 2017 and will take a similar amount in 2018. The group production of DAODs for data and MC took about 15% of the resources while a smaller fraction went into the reprocessing of the 2015-2017 data and distributed analysis by individual ATLAS users. MC simulation, reconstruction, derivation production and data reprocessing use the multi-processing workflow to allow efficient memory usage of the available Grid resources with a maximum of around 2 GB per CPU core.

In 2017 almost 10 billion events have been simulated using full simulation and about 1.6 billion events using fast simulation. Fast simulation uses a parametrised simulation of the ATLAS calorimeter and allows to simulate events with less CPU time used but with a slightly less precision. Fig. 6 right shows the breakdown of produced full simulation MC events per resource type. While Grid resources provide the vast majority of the capacity available to ATLAS, the contribution of Clouds and HPCs (14% and 9% shares respectively of full MC simulation event production) is significant. HPC resources are often limited in their capability to a few workflows, contributing mostly to MC simulation and, in the case of architectures different than x86_64, to event generation only.

The 'Data Processing' category had two major phases of activity with the reprocessing of 2016 data in the spring 2017 and the reprocessing of 2017 data in Winter 2017/18. A major campaign of derivation production using all 2015-2017 data and some MC was done in Winter 2017/18.

5 Tier-0

The Tier-0 centre at CERN is responsible for the safe-keeping of the RAW data and its distribution to the Tier-1s. The first pass calibration and reconstruction is done within days after the data has been taken. The ATLAS Tier-0 delivered excellent throughput during 2017 and

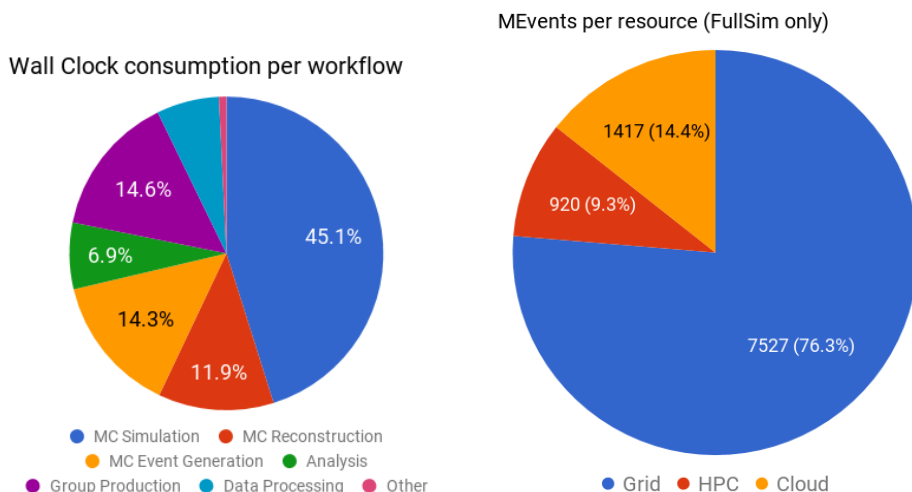


Figure 6. Left: CPU wall clock consumption per different workflow activities in 2017. Right: Number of simulated events split by resource types in 2017 where "MEvents" corresponds to million events.

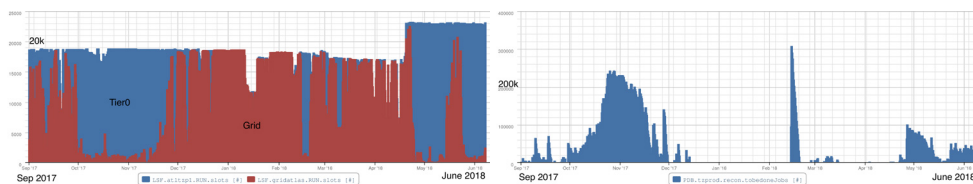


Figure 7. Left: Number of jobs running at the Tier-0 at CERN in the past 9 months with the first pass reconstruction (blue) and the Grid processing (red). Right: Number of queued jobs at the Tier-0.

2018 so far, benefiting from the workflow and infrastructure improvements. The CPUs cluster is configured without hardware hyper-threading, providing 4 GB of memory available for each slot and making it possible to run the full Tier-0 processing chain in single core mode. The workflow reads the RAW detector data and produces all output formats in a single step, keeping the intermediate ones in memory and so reducing the IO intensity. Fig. 7 left shows the number of running jobs in the LSF batch system dedicated to the ATLAS Tier-0. The LHC operation efficiency reached 74%, during a week in October 2017, with an average between 50% and 60% during the data taking period. In the periods of inactivity for first pass reconstruction (in blue), the system is backfilled with Grid processing (in red), with the Tier-0 resources constantly in full operation. It can be seen that in the period from mid-October to the end of November, when the LHC was operating in a mode of higher average number of proton-proton interactions, the Tier-0 was fully occupied by prompt reconstruction jobs and a backlog of up to 250k jobs was accumulated (Fig. 7 right). This was due to a higher average trigger rate (up to 1.5 kHz) and longer time/event needed to reconstruct events at higher pileup values. During the LHC winter shutdown the Tier-0 was used mainly for Grid activities and in Spring 2018 switched back to Tier-0 prompt reconstruction workflows after the LHC restart.

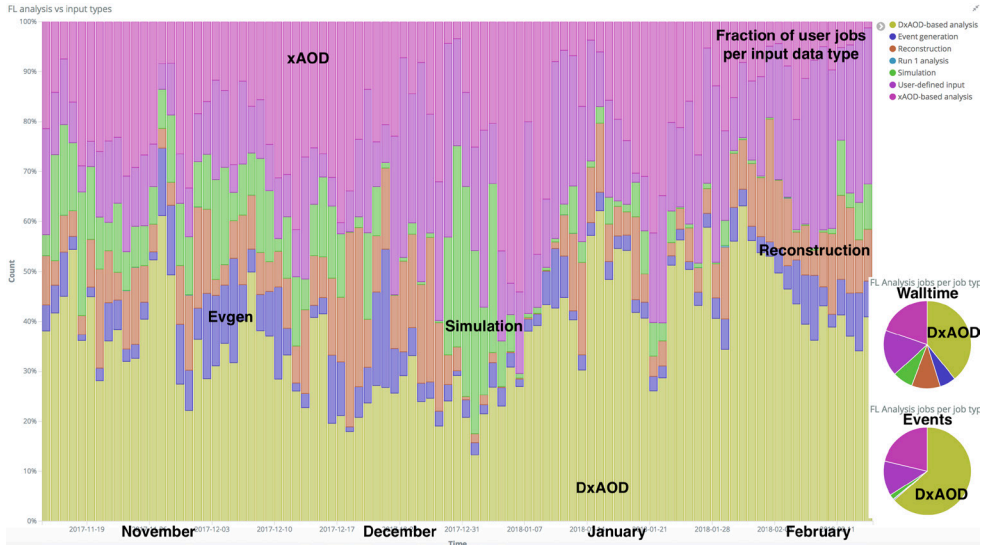


Figure 8. Distributed analysis activity for November 2017 - February 2018 split by the different input files types and workflow types. "xAOD" and "DxAOD" refer to AOD and DAOD output files of the reconstruction and derivations workflows as explained in the text. "Evgen" denotes the event generation.

6 Distributed Analysis

ATLAS distributed analysis is performed on the Grid by several hundred individual ATLAS users. A large fraction of events are processed from DAODs (see. Fig. 8). Some CPU time is devoted to process or generate other formats mainly for developing or testing new software algorithms or for systematic studies of physics analysis.

7 Monitoring and Analytics

The various components of the ATLAS distributed computing system are monitored in different ways using multiple tools. Usually there is dedicated monitoring for each system. An example of such a single system monitoring is shown in Fig. 9 left with a screen shot of the PanDA workflow management system monitor webpage. This monitor is based on the PanDA job information stored in the Oracle SQL database and provides extensive and detailed information about the system.

Since ATLAS distributed computing needs to account, monitor and optimise the usage of all computing services and resources combined, an analytics platform has been developed and setup that collects information from all components and combines them where possible. Fig. 9 right shows a screen shot of a Kibana dashboard of new ATLAS analytics platform [8]. This system is based on an NoSQL elasticsearch cluster [9], custom made data collectors and data enrichment on a Kubernetes cluster [10]. Many detailed studies of the ATLAS Grid system and workflow performance have been carried out and several workflow improvements have been introduced using this combined information [11].

8 Evolution of the systems

The ATLAS workflow and data management systems and the Grid infrastructure are still evolving. Some major new features are being developed or commissioned.

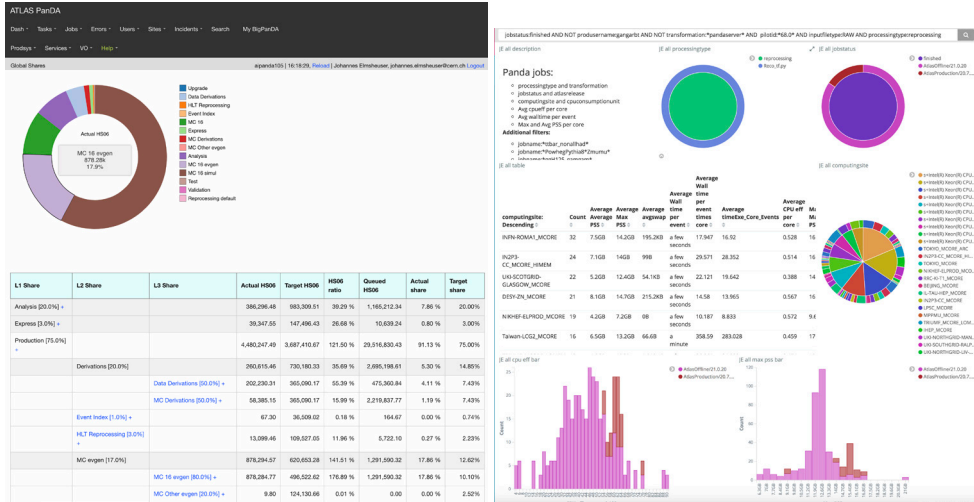


Figure 9. Left: Screenshot of the PanDA monitoring webpage with a breakdown of the global shares of the different activities. Right: Example of an analytics dashboard using Kibana.

To exploit all available resources an efficient job scheduling is necessary. The rather diverse resource usage capability of the ATLAS workflow management will be improved with developments of Harvester [12], Event Service and Event Streaming Service [13] and a new version of the PanDA pilot as described further in this section.

ATLAS is running a diverse mixture of workflows and different processing campaigns at different times. Single and multi process workflows are submitted to many resources. So called global fair-shares and unified queues have been introduced to limit the CPU slots per activity and boost activities when needed. The currently partitioned production and analysis queues will be unified into a single queue per site to allow more efficient scheduling. The PanDA pilot code that executes the job payload on the site worker nodes has been redesigned for easier site support and is currently being commissioned.

Harvester is being developed to optimise the allocation of jobs to CPUs from different resources like the Grid, HPC and Cloud. Since every HPCs has its own job scheduling mechanism and rather restricted user access policies, Harvester was developed initially for HPCs as a common transparent interface to these resources. Now it also allows access to Grid and Cloud resources in the so-called PanDA pilot pull and push modes and is currently being commissioned at many ATLAS WLCG Grid sites.

The development of the Event Service (ES) workflow and Event Streaming Service (ESS) allows fine grained event level processing instead of the classical file level based processing. The ES assigns fine-grained processing jobs to workers and streams out the data in quasi-real time, ensuring fully efficient utilisation of all resources, including volatile resources. The ESS is in development to asynchronously deliver only the input data required for processing when it is needed, protecting the application payload from WAN latency without creating expensive long-term data file replicas.

The data management system Rucio is very successfully handling the large amounts of ATLAS data. This has triggered interest in Rucio by other HEP experiments and communities. Several new features are being developed for Rucio: the Rucio input file mover has been added to the PanDA pilot code, the awareness of site file caches is being added, object stores for input and output files are supported and the automatic creation of archive files has

been added. In addition the activities within the WLCG DOMA (data organisation, management, access) project [14] to explore possibilities to overcome disk shortage in HL-LHC are ramping up and ATLAS is actively participating. A very successful R&D project with Google has been started which allows the integration of Rucio and distributed analysis using Harvester [15].

The site infrastructure is evolving as well and the support for containers using Singularity [16] and Docker [17] is under development at the sites and within PanDA jobs for production and user analysis usage.

9 Conclusions

The ATLAS workflow and data management systems are robust, performant and can easily cope with the higher Run 2 LHC performance. The processing power, storage capacities and network are used to their limits. There are presently no scaling issues and each subsystem is able to sustain the large loads. New workflows and technologies are being integrated but the slight shortage of storage has to be addressed in the future.

References

- [1] ATLAS Collaboration, JINST **3** S08003 (2008).
- [2] ATLAS luminosity measurements in Run 2, URL <https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2> [accessed 2018-09-06]
- [3] T. Maeno et al. [ATLAS Collaboration], J. Phys. Conf. Ser. **331** 072024 (2011).
- [4] V. Garonne et al. [ATLAS Collaboration], J. Phys. Conf. Ser. **898** no.6, 062019 (2017).
- [5] Worldwide LHC Computing Grid project, URL <http://cern.ch/lcg> [accessed 2018-09-06]
- [6] W. Wu, D. Cameron, A. Filipcic, PoS ISGC 2017 014 (2017).
- [7] File Transfer Service (FTS), Version 3.7.10 available from <https://fts.web.cern.ch> [accessed 2018-09-06]
- [8] I. Vukotic et al., *ATLAS Analytics and Machine Learning Platforms*, These Proceedings (2019).
- [9] Elasticsearch, Version 6.4.0 available from <https://www.elastic.co/downloads/elasticsearch> [accessed 2018-09-06]
- [10] Kubernetes, Version 1.11 available from <https://kubernetes.io/docs/setup/release/notes/> [accessed 2018-09-06]
- [11] J. Elmsheuser et al., *ATLAS Grid Workflow Performance Optimization*, These Proceedings (2019).
- [12] T. Maeno et al., *Harvester : an edge service harvesting heterogeneous resources for ATLAS*, These Proceedings (2019).
- [13] N. Magini et al., *Towards an Event Streaming Service for ATLAS data processing*, These Proceedings (2019).
- [14] S. Campana, *Architecture and prototype of a WLCG data lake for HL-LHC*, These Proceedings (2019).
- [15] M. Lassnig et al., *The ATLAS & Google "Data Ocean" Project for the HL-LHC era*, These Proceedings (2019).
- [16] Singularity, Version 2.52 available from <https://singularity.lbl.gov/all-releases> [accessed 2018-09-06]
- [17] Docker, Version 18.06 available from <https://www.docker.com/products/docker-desktop> [accessed 2018-09-06]